Modernizing the Mathematics Curriculum – Part 2

The Case of Implicit Differentiation:

An Opportunity to Make Sense or Just Follow a Prescribed Technique?

By Scott Adamson (s.adamson@cgc.edu)

Introduction

Try searching the internet for videos related to implicit differentiation. Mostly, what you will find are videos showing how to *perform the technique* of implicit differentiation. What is missing, in my opinion, is a discussion of how the method works, why we do what we do in performing the technique, as well as an explanation of how the results make sense.

When discussions about modernizing the curriculum come up, they often involve two areas: what *topics* we teach and what we want *students to learn* about these topics. In Part 2 of this short series on modernizing the curriculum, we focus on what we want students to learn even when choosing to teach a common topic in the Calculus 1 curriculum: implicit differentiation.

We begin by defining what we mean by an implicitly defined function and contrasting it with the more familiar explicitly defined function. Consider the table:

Explicitly Defined Function	***Implicit Functions***
$y = 4x - 6$	$y - 4x = 6$
$y = 2x^2 - 3x + 4$	$x^2 + 2y^2 = xy + y^3$
$y = \sin(2x) + 4$	$2x - y^2 = \sin(xy)$
$y = 2e^{3x}$	$e^{xy} + 2 = 2xy^2 - 5y$

We say that the examples on the left side of the table are relationships between x and y such that y is explicitly defined to be in terms of x.

The examples on the right side of the table are relationships between x and y such that y represents a function(s) producing corresponding values of y for given input values x that satisfy the equation.

Graphs of implicitly defined curves have rates of change that can be computed with the idea of the derivative. The technique or process that is typically explained in a standard Calculus 1 course known as implicit differentiation, followed by questions that are often not considered, is outlined below.

1. Start with the implicitly defined equation and take the derivative of both sides with respect to x.
   Why can we take the derivative of both sides of an equation? Can we always take the derivative of both sides of any equation?

2. When taking the derivative of anything involving *y*, think of *y* as a function of *x* and take the derivative accordingly.
   Why do we do this and how does it make sense? Specifically, what is the function of x referenced in this step and why do we have to represent y as some mysterious function of x?
3. Solve the resulting equation for y' or $\frac{dy}{dx}$ depending on preference.

   Why is algebra so hard?
4. Write your answer as $y' =$ or $\frac{dy}{dx} =$

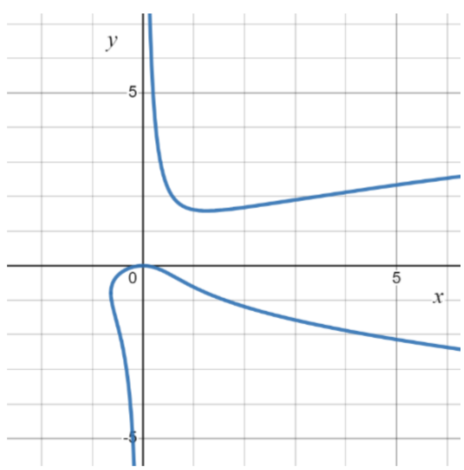
   What does this answer mean? If we were trying to find the rate of change of the curve, why does the answer indicate that we found the derivative of y? What is y and how is it connected to f(x)?

One additional note: teachers often discuss and debate whether to use the notation y' or $\frac{dy}{dx}$ when teaching the technique of implicit differentiation. The discussion often revolves around which notation students like best or which is easier for students to wade through the symbolic procedure and successfully produce the right answer. One goal of this paper is to push the professional discussions toward the mathematical meanings and conceptual understandings entailed in the technique of implicit differentiation rather than to be solely focused on a "best" procedure to help students obtain the correct answer.

Part 1 – Making Sense of Implicitly Defined Functions

In an attempt to develop a more robust understanding of the technique of implicit differentiation, consider the following implicitly defined equation and its graph:

$$x^2 + y = xy^2$$

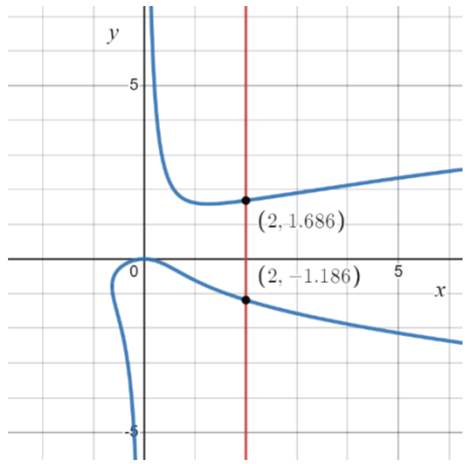


We can extend our understanding of graphs as containing infinitely points that satisfy, in this case, the condition that $x^2 + y = xy^2$ and imagine different inputs, *x*, and corresponding values of *y* that satisfy the equation.

For example, when $x = 2$, we need to think about a corresponding value of y such that

$$2^2 + y = 2y^2$$

By looking that the graph, we see that there are two values of y that satisfy this condition (note that these values are rounded to the nearest thousandths place).



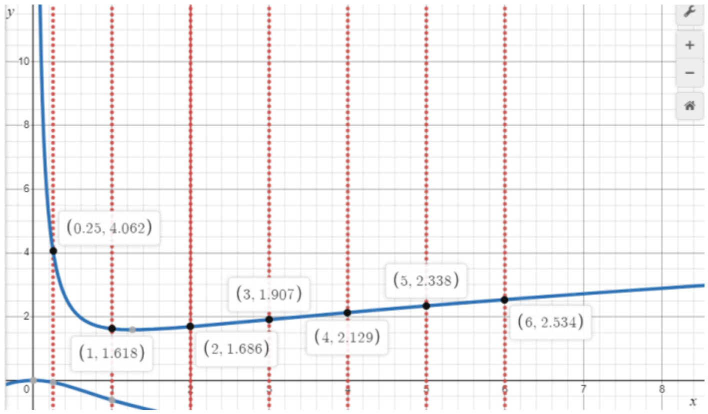
$$2^2 + 1.686 = 2(1.686)^2 \qquad\qquad 2^2 + (-1.186) = 2(-1.186)^2$$
$$5.686 = 5.686 \qquad\qquad 2.814 = 2.814$$

(recall that there is rounding error)

Imagine thinking about this process of determining corresponding y_1 values for many, many input values, x_1. For the sake of this explanation, we will first consider the following table of input values, x_1, and determine their corresponding *positive* values, y_1, such that $x^2 + y = xy^2$. We will use Desmos to find these values as shown in the series of images that follow.

x_1	y_1
0.25	------
1	------
2	------
3	------
4	------
5	------
6	------

By graphing $x = 0.25, 1, 2, 3, 4, 5, 6$, we can determine the intersection points of these vertical lines with the graph of $x^2 + y = xy^2$ and thus determine the corresponding positive values of y_1 for each input value, x_1 (as shown).



x_1	y_1
0.25	4.062
1	1.618
2	1.656
3	1.907
4	2.129
5	2.338
6	2.534

Note that we can assert that there is a single positive value of y_1 that corresponds with the given input value of x_1 that satisfies $x^2 + y = xy^2$. In other words, y_1 is a function of x_1 or $y_1 = f(x_1)$. Since we can say that the positive values of y that satisfy the equation $x^2 + y = xy^2$ are a function of the input values, x, it follows that:

$$x^2 + f(x) = x\left(f(x)\right)^2$$

In this context, we are using the table (see below) to represent the function $f(x)$ and we acknowledge that the table only shows certain input values, x, and many, many more could be determined. Nonetheless, the table is a representation of the function $f(x)$.

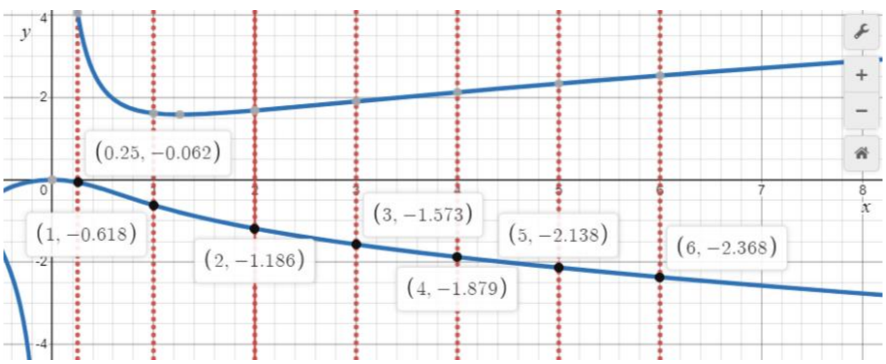
x_1	y_1
0.25	4.062
1	1.618
2	1.656
3	1.907
4	2.129
5	2.338
6	2.534

This table is a representation of a function $y = f(x)$.

Similarly, we can determine corresponding *negative* values of y_2 that satisfy the equations for the selected input values of x_2. We will use Desmos to find these values as shown in the series of images that follow.

x_2	y_2
0.25	------
1	------
2	------
3	------
4	------
5	------
6	------

By graphing x = 0.25, 1, 2, 3, 4, 5, 6, we can determine the intersection points of these vertical lines with the graph of $x^2 + y = xy^2$ and thus determine the corresponding negative values of y_2 for each input value, x_2 (as shown).



x_2	y_2
0.25	−0.062
1	−0.618
2	−1.186
3	−1.573
4	−1.879
5	−2.138
6	−2.368

Note that we can assert that there is a single negative value of y_2 that corresponds with the given input value of x_2 that satisfies $x^2 + y = xy^2$. In other words, y_2 is a function of x_2 or $y_2 = f(x_2)$. Since the negative values of y that satisfy the equation $x^2 + y = xy^2$ are a function of the input values, x, it follows (once again) that:

$$x^2 + f(x) = x\left(f(x)\right)^2$$

In this context, we are using the table (see below) to represent the function $f(x)$ and we acknowledge that the table only shows certain input values, x, and many, many more could be determined. Nonetheless, the table is a representation of the function $f(x)$.

x_2	y_2
0.25	−0.062
1	−0.618
2	−1.186
3	−1.573
4	−1.879
5	−2.138
6	−2.368

This table is a representation of a function $y = f(x)$.

We now must acknowledge and embrace an important idea. When we think about $x^2 + y = xy^2$, we must consider that for any input value x, the corresponding value of y that satisfies the equation is at least one function of x. As seen above, for any input value x, there is a positive and a negative corresponding value of y that satisfies the equation. That is, we are using $f(x)$ to represent these functions and we express the equation as:

$$x^2 + f(x) = x\left(f(x)\right)^2$$

It is the case that the function(s) $f(x)$ are the corresponding y values for any input values x that make the equation true. Therefore, finding the derivative of $f(x)$, $\frac{d}{dx}\left[f(x)\right]$, will be equivalent to finding the rate of change of the implicit curve at the indicated point on the curve, (x, y). How can we find the rate of change of the curve? That is, how can we compute $\frac{d}{dx}\left[f(x)\right]$?

Part 2 – Differentiating Both Sides?

It is common for calculus teachers to teach student to "take the derivative of both sides of the equation with respect to x" when showing the technique of implicit differentiation.

But what is the mathematical reasoning that allows this to happen? Can we "take the derivative of both sides" of any equation?

Consider the equation

$$2x + 4 = 3$$

What if we took the derivative of both sides of this equation with respect to x?

$$\frac{d}{dx}(2x+4) \stackrel{?}{=} \frac{d}{dx}(3)$$

$$2 \stackrel{?}{=} 0$$

Clearly, we have created a contradiction which means that there must be more to the story when we "take the derivative of both sides" when engaged in implicit differentiation!

Rather than "taking the derivative" of both sides of an equation, consider the following reasoning in the context of the equation analyzed in Part 1: $x^2 + y = xy^2$ or $x^2 + f(x) = x\left(f(x)\right)^2$.

Let's define two new functions $m(x)$ and $n(x)$ where one represents the left side of the equation and one the right side of the equation $x^2 + f(x) = x\left(f(x)\right)^2$.

$$m(x) = x^2 + f(x)$$

$$n(x) = x\left(f(x)\right)^2$$

Therefore, we can assert the following:

$$x^2 + f(x) = x\left(f(x)\right)^2$$
$$m(x) = n(x)$$

It is true that if $m(x) = n(x)$, it follows that $\frac{d}{dx}\left[m(x)\right] = \frac{d}{dx}\left[n(x)\right]$. That is, if two functions are equivalent as in $m(x) = n(x)$, then it must be true that their derivatives are also equivalent, $\frac{d}{dx}\left[m(x)\right] = \frac{d}{dx}\left[n(x)\right]$. So, while "taking the derivative" of both sides of the equation $x^2 + f(x) = x\left(f(x)\right)^2$ will ultimately produce the correct result, the reasoning behind the move is faulty since "taking the derivative" of both sides of any equation has been shown to produce false results in some cases (recall the previous example where differentiating both sides of an equation resulted in the false statement 2 = 0). Rather, we are taking the derivatives of two equivalent functions and creating another true statement, $\frac{d}{dx}\left[m(x)\right] = \frac{d}{dx}\left[n(x)\right]$.

Let's apply this way of thinking to the current situation: $x^2 + f(x) = x\left(f(x)\right)^2$. Note the need for the product rule and chain rule in the appropriate places.

$$x^2 + f(x) = x\left(f(x)\right)^2$$
$$m(x) = n(x)$$
$$\frac{d}{dx}\left[m(x)\right] = \frac{d}{dx}\left[n(x)\right]$$
$$2x + \frac{d}{dx}\left[f(x)\right] = x \cdot 2f(x) \cdot \frac{d}{dx}\left[f(x)\right] + 1 \cdot (f(x))^2$$

Recall that it is our goal to find the derivative of $f(x)$, $\frac{d}{dx}\left[f(x)\right]$, which is equivalent to finding the rate of change of the implicit curve at any point on the curve, (x, y). So, our goal is to now solve for $\frac{d}{dx}\left[f(x)\right]$ as shown.

$$\begin{aligned}
2x+\frac{d}{dx}\left[f(x)\right] &= x\cdot 2f(x)\cdot\frac{d}{dx}\left[f(x)\right]+1\cdot(f(x))^2 \\
2x+\frac{d}{dx}\left[f(x)\right] &= 2x\cdot f(x)\cdot\frac{d}{dx}\left[f(x)\right]+(f(x))^2 \\
\frac{d}{dx}\left[f(x)\right]-2x\cdot f(x)\cdot\frac{d}{dx}\left[f(x)\right] &= (f(x))^2-2x \\
\frac{d}{dx}\left[f(x)\right]\left(1-2x\cdot f(x)\right) &= (f(x))^2-2x \\
\frac{d}{dx}\left[f(x)\right] &= \frac{(f(x))^2-2x}{1-2x\cdot f(x)}
\end{aligned}$$

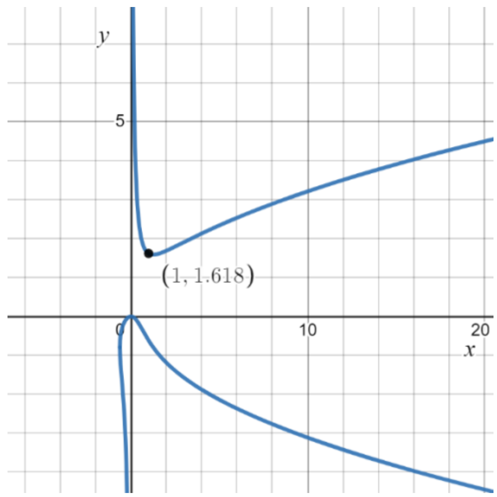
Note that the derivative of $f(x)$ is determined by both the input value x and the corresponding value of y, $f(x)$. Equivalently,

$$\begin{aligned}
\frac{d}{dx}\left[f(x)\right] &= \frac{(f(x))^2-2x}{1-2x\cdot f(x)} \\
\frac{d}{dx}\left[y\right] &= \frac{y^2-2x}{1-2x\cdot y} \\
\frac{dy}{dx} &= \frac{y^2-2x}{1-2x\cdot y}
\end{aligned}$$

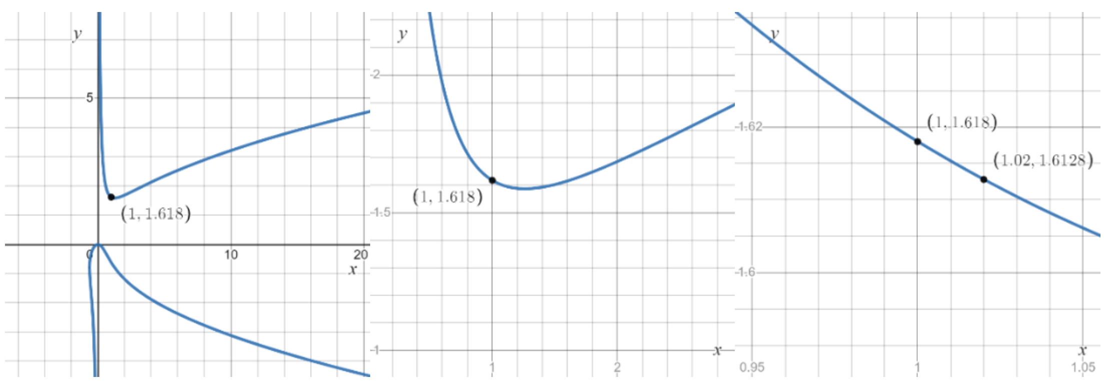
Recall that we acknowledged and embraced the idea that the function $f(x)$ might represent more than just one function as in the example in Part 1. Now we can see how this is allowable. In the example in Part 1, we had one function that produced corresponding *positive* y values for given input values, x. We had another function that produced corresponding *negative* y values for given input values, x. Now we see that the derivative, $\frac{d}{dx}\left[f(x)\right]$, depends on the point (x,y) in question. For positive values of y, the derivative will represent the rate of change of the implicitly defined curve in Quadrant 1 (where y is positive). For negative values of y, the derivative will represent the rate of change of the implicitly defined curve in Quadrant 4 (where y is negative). That is, the derivative $\frac{dy}{dx}=\frac{y^2-2x}{1-2x\cdot y}$ will be computed correctly for the given ordered pair (x,y) in question.

Part 3 – Making Sense of the Results

Let's consider the graph of the equation $x^2 + y = xy^2$ once again and specifically focus on the point (1, 1.618). Think about the rate of change of this implicitly defined curve at this point. The rate of change should be negative.



If we zoom in on the point (1, 1.618), we can visualize the local linearization of the curve at this point and estimate its rate of change. Analyze the series of images to represent the zooming in process.



By using the final image, we can estimate the rate of change of the curve, $\frac{d}{dx}\left[f(x)\right]$, at the point (1,1.618) by computing the rate of change using the points (1,1.618) and (1.02,1.6128)

$$\begin{aligned}\frac{d}{dx}\left[f(x)\right] &\approx \frac{1.6128 - 1.618}{1.02 - 1} \\ &\approx \frac{-0.0052}{0.02} \\ &\approx -0.26\end{aligned}$$

Now, let's use the derivative, $\frac{d}{dx}\left[f(x)\right]$, that we computed using the process of implicit differentiation earlier. We should arrive at a result close to this result of –0.26. We will compute $\frac{dy}{dx}=\frac{y^2-2x}{1-2x\cdot y}$ at the point (1,1.618).

$$\begin{aligned}\frac{dy}{dx}&=\frac{y^2-2x}{1-2x\cdot y}\\ \left.\frac{dy}{dx}\right|_{(1,1.618)}&=\frac{(1.618)^2-2\cdot 1}{1-2\cdot 1\cdot 1.618}\\ &=\frac{0.617924}{-2.236}\\ &\approx -0.276\end{aligned}$$

Our results of –0.276 confirm the results from estimating the rate of change of the curve previously resulting in a rate of change of –0.26.

Conclusion

One might argue that the idea presented in this paper could make the teaching of the topic of implicit differentiation more complicated and messier for our Calculus 1 students. For example, it is true that manipulating the $\frac{d}{dx}\left[f(x)\right]$ symbols might be messier than manipulating a y' or even a $\frac{dy}{dx}$ symbol. However, the focus of our teaching might be better placed on the important ideas involved in the topic rather than focusing on ways to make the symbolic manipulations easier for students. If getting the right answer was the only or primary goal of implicit differentiation, then one could learn to use Geogebra to produce right answers!

To conclude this essay, we leave you with two highlights that focus on important ideas related to implicit differentiation:

Highlight 1: What is the mathematical foundation that allows taking the derivative of both sides of an equation? It is not an algebraic move. Rather, it is about making truth claims. If two functions are equivalent, $m(x)=n(x)$ (a truth claim), then their derivatives must be equivalent as well: $\frac{d}{dx}\left[m(x)\right]=\frac{d}{dx}\left[n(x)\right]$ (another truth claim).

Highlight 2: In an implicitly defined function, the "*y*" value is a function (or more than one function) of *x*. That is, for input values *x*, the corresponding values of *y* that satisfy the implicit equation are a function(s) of *x*. So, when computing $\frac{d}{dx}\left[f(x)\right]$, we are therefore also computing

$\frac{dy}{dx}$! That is, we are indeed finding the rate of change of the implicitly defined curve at any point (x, y)!